开展全科教育干预研究：从设计到发表

10.12114/j.issn.1007-9572.2022.0597

邹川[1]、欧健铭[1]、曾欣[1]、彭涛[1]、周燕[1]、肖纯瑶[1]、陶洪霞[2]、陈庆奇[3]、林恺[4*]

基金项目：四川省医学（青年创新）科研课题（项目编号：Q20100）；**成都市医学科研课题**（项目编号：2021062）；**院级研究基金**（项目编号：JGZX202215）；**广东省本科高校临床教学基地教学改革项目**（项目编号：2019JD108）

1、611130 四川省成都市，成都市第五人民医院 全科医学科

2、611137 四川省成都市，成都中医药大学 医学与生命科学学院

3、518000 广东省深圳市，联合医务集团

4、515041 广东省汕头市，汕头大学医学院第一附属医院 全科医学科

*通讯作者：林恺，主治医师，讲师；E-mail：Klin1@our.ecu.edu.au

【摘要】随着中国基层医疗的发展，近 10 年中国将培养近 40 万全科医生，其中培养具有胜任力的全科医生至关重要，这需要全科教育者探索大量适合中国且基于证据的全科教育理念、方法和内容。目前，国内全科教育干预研究的质量较差，大部分全科教育师资和研究者尚缺乏教育干预研究的相关培训。本文旨在介绍全科教育干预研究：从设计到发表的流程，将研究分为四个阶段，共 13 个步骤。四个阶段分别为研究准备（产生研究灵感、文献综述、融入理论/概念框架、精炼研究问题）、研究设计（试验设计、干预措施、结果评估）、研究执行（建立研究团队、获取研究资源、申请研究伦理、项目实施）、发表和评价阶段（文章撰写、反思与评价）。以期为全科教育研究者与实践者开展教育干预研究提供研究方法和思路，产生高质量的教育研究“证据”，进一步改善全科教育培训质量，培养具有胜任力的全科医生。

【关键词】全科医学；教育研究；干预性研究

1. Department of General Practice, Chengdu Fifth People’s Hospital, Chengdu,611130, China

2. College of Medicine and Life Sciences, Chengdu University of Traditional Chinese Medicine, Chengdu 611137, China

3. Department of General Practice, Shenzhen 518000, China, UMP Healthcare China

4. Department of General Practice, The First Affiliated Hospital of Shantou University Medical College, Shantou 515041, China

**Corresponding authors: Lin Kai, Attending Physician, Lecturer; E-mail: Klin1@our.ecu.edu.au*

【Abstract】With the development of primary care in China, nearly 400,000 general practitioners will be trained in China in next 10 years, which requires a large number of China-appropriate and evidence-based general practice education concepts, methods, and contents. At present in China and abroad, the quality of educational intervention studies is poor, and most faculty and researchers in general practice lack training in educational intervention studies. The purpose of this paper is to introduce the process of educational intervention research in general practice from design to publication, dividing the research into four stages with 13 steps. The four stages are research preparation (generating research inspiration , literature reviews, incorporating a theoretical/cognitive framework, refine research question), research design (trial design, intervention, outcome evaluation), research implementation (establishing a research team, obtaining

research resources, applying for research ethics, program execution), and publication and evaluation phases (article writing, reflection and evaluation). The aim is to provide research methods and ideas for educational researchers and practitioners in general practice to conduct educational intervention studies, generate high-quality educational research "evidence", further improving education , and train competent general practitioners.

【Keyword】General Practice; Educational Research; Intervention

为加强基层医疗卫生服务体系建设，推动分级诊疗，培养具有胜任力的全科医生显得至关重要。2018 年国务院办公厅《关于改革完善全科医生培养与使用激励机制的意见》指出，到 2030 年城乡每万名居民拥有 5 名合格的全科医生，意味着近 10 年中国将培养 40 万全科医生[1]。然而，中国全科医学起步较晚，所采用的全科教育培训体系、理念和方法大多借鉴国际经验，再加上中国独特的卫生服务体系和文化传统，培训的有效性和适用性不得而知。

医学教育已从基于理论的教学逐渐演变为基于证据的教学，教师在进行教学实践时，需要充分考虑培训方法的证据水平，为学生提供最佳证据的医学教育[2]。而医学教育干预（Educational Intervention）研究能为教师的教学实践提供可靠的判断依据。教育干预主要是指采用新的培养计划、课程模式、教学手段，在教育教学活动中有目的、有计划、有针对性地进行教育实践，从而改革旧的教育体系或做法，提高教学效果[3]。随着全科医学教育领域的迅猛发展，一方面需要吸取国际全科医生的培养经验，另一方面有赖于在国内开展高质量的全科教育干预研究，产生更多的教育研究“证据”，做到循证教育实践。

然而，许多研究者会将医学教育研究视为是一门“软”科学，因此在研究设计时缺乏足够的严谨性和科学性。一篇医学教育研究的系统综述显示其纳入研究总体质量较差，如没有明确的方法学描述，缺乏对照组或机构审查委员会的批准[4]。另一篇全科医学教育研究的范围综述也证实大部分研究需要在方法学上进行优化，其中 23%的定量研究采用前后测验，近一半采用问卷调查作为主要的数据收集方法，只有 10%采用随机对照试验[5]。研究设计的不足会降低研究的效度，进而限制教育成果的推广。

本文旨在通过文献综述和笔者的实践经验，介绍全科教育干预研究从设计到发表的四个阶段，13 个步骤，为全科教育研究者与实践者开展教育干预研究提供研究方法和思路。

1　阶段一：研究准备

1.1　步骤 1、产生研究灵感

产生研究灵感是构建研究问题的第一步。一方面，全科教育研究者可以通过阅读大量全科或教育类的相关文献和参与学术会议获得研究灵感[6]。另一方面，灵感来源的最重要的途径是全科教育者的实践经历[7, 8]。全科教育者和实践者会持续遇到各种临床问题和教学困境，也会接受各种有助于培养全科学员的新理念和方法，如叙事医学、共同决策、情景模拟等，这些都可以为研究提供有价值的方向。

1.2　步骤 2、文献综述

当产生研究灵感或研究问题后，文献综述可以帮助研究者对感兴趣的方向的研究现状建立一个全面、清晰的图景，并发现研究中存在的空白，也避免重复他人的研究。同时，研究者可以在文献综述的过程中吸取和借鉴其他学者的研究经验，如试验设计、干预方法、评估工具、数据分析等[9]，通过学习和比较进一步构思自己的研究计划。

1.3　步骤 3、融入理论（Theory）/概念框架（conceptual framework）

教育干预研究中的理论/概念框架好比药理的机制。研究者除了要明确干预是否产生效果，还要通过各类理论/概念框架指导干预的设计，并且在研究发表时阐明产生效果的原因[10]，这也是教育研究与临床研究最为显著的区别。在以往的全科医学教育研究和论著中，往往存在理论/概念框架的缺失[11]。2007 年及 2019 年两篇对主要医学教育期刊发表文献的综述表明，近一半论文缺少明确的理论或概念框架支持[5, 12]。这一问题严重影响了全科医学领域教育研究设计的严谨性，桎梏了研究成果的推广和应用。

Bordage 等人将理论/概念框架描述为“思考问题或研究的方式，或呈现事物间复杂联系的方法”，主要包括一系列基于观察或实证研究而发展起来的教育理论（如刻意练习）、模型（如库伯学习模型）或实践原则[13]。在研究中使用理论/概念框架可以帮助阐明问题的本质，并指导发展可能的解决方案或干预措施，并帮助同行或读者快速理解研究基础和假设，更利于研究成果的推广应用[14]。在教育研究中，理论/概念框架就像是不同的“聚光灯”，反映和强调研究问题的某个方面，只是对客观现实的局部呈现[13]。例如，一个研究的目标是“提升全科医生医患沟通技能”，认知行为主义强调心理策略和实践评估，建构主义强调知识和技能获取的步骤，而社会学习理论关注师生或学生间的互动[15]。因此，设计良好的教育干预研究往往基于概念框架发展以提高研究的严密性[6]。设计时，既可以使用单一的概念框架，也可以尝试运用多种概念框架，从而达到更理想的干预效果。本文列举了全科教育研究中的三个常用概念框架，供研究者理解和参考（见表 1）。

表 1　全科医学教育干预研究中的理论/概念框架举例

Table 1 Examples of theory/conceptual frameworks of medical educational intervention studies in general practice

理论/概念框架	理论内容	全科医学领域的应用
反思性学习理论（Reflective learning）	反思性学习是通过对学习活动过程的反思，来强化学习效果。 反思是对自己的思维过程、思维结果进行再认识的检验过程。学习中的反思如同生物体消化食物和吸收养分一样，是无法代替的。当代建构主义学说认为：学习要在活动中进行建构，要求学生对自己的活动过程不断地进行反省、概括和抽象。	Shaughnessy 等人在家庭医学住院医生培训项目中引入了持续的反思练习。他们组织住院医生进行每周三次，针对临床实践的反思和记录，并通过焦点小组进行质性研究，旨在增强住院医生的自我发展能力。[16]
实践共同体（Community of practice）	知识、共同体和实践是实践共同体的三大要素。实践共同体的特点是，共同体成员有着近似的身份或兴趣点，并在共同的目标引导下一起学习、应用和创造知识。在实践学习中，共同体分享着共同信念和理解，并且通过沟通协商进行实践活动；最终形成交流合作、共享资源、互相帮助、获得知识，然后又有能力将知识运用到实践中。	Malaty 等人让家庭医学住院医生组成临床实践团队，目的是学习临床实践的管理和利用患者的临床数据提高诊疗质量。研究者定期向团队提供临床实践的管理数据，并开展每季度一次的教学研讨，由医疗主任和住院医师教员提供关于如何通过机构流程改善此类指标的教育。最终评价学员各方面知识和能力。[17]
成人学习理论（Adult learning）	成人学习理论的提出从成人视角研究了学习者的心理特征和行为需要。成人学习者已有经验为基础，且具有清晰自我概念，因此其学习以解决问题为中心、以主观需求为导向、受内在动机驱动。我国近年来在医学教育领域运用该理论加强转化学习理论的研究与应用，旨在通过学习以帮助学生转化自己。在成人学习理论视角下，更重视对学生自主学习和批判性思维能力与习惯的培养。	基于成人学习理论，加拿大阿尔伯塔大学设计了为期 2 年的家庭医学住院医师培训计划。其中，轮转课程涵盖了 23 个家庭医学核心主题；学习过程还包含一系列小组活动，如：基于问题的学习模块，日记俱乐部和考试准备会议等。最后，通过多源方法评估整个教学培训。[18]

教育干预研究中的理论/概念框架通常来源于教育学、心理学甚至社会学，瀚如烟海且晦涩难懂的理论或模型往往会让研究新手们望而生畏。我们建议在选择概念框架前，对相关主题进行系统检索，了解并学习既往研究所使用的概念框架。当然，与经验丰富的教育研究者合作会是一条捷径。此外，推荐阅读《医学教育研究概论》[19]，书中对教育研究的各种常用理论模型进行了非常详细的论述。

1.4　步骤 4、精炼研究问题

研究灵感转化为好的教育干预研究问题很具有挑战性，可以尝试运用循证医学中的 PICOT 格式，帮助全科教育研究者进一步提出更清晰和细化的研究问题。PICOT 包括人群（Population）、干预组（Intervention）、对照组（Comparision）、结果（Outcomes）和时间(Timing)[20]。以我们的研究为例，研究问题是“与沟通理论课程（C）相比，基于卡尔加里指南的沟通培训沙龙（I），是否可以提升二年级全科住院医师（P）在 3 月后（T）的门诊接诊沟通能力（O）”。FINER 标准常用来评估构建的研究问题质量，包括可行性（Feasible），有趣性（Interesting），创新性（Novel），伦理性（Ethically）

和相关性（Relevant）[21]。其中，相关性意味着研究问题应与全科教学实践密切相关，旨在解决当前全科教育存在的困难和挑战。通过上述的两个方法，逐渐让研究灵感精炼为具体的、可以回答并有实践创新价值的教育研究问题。

2　阶段二：研究设计

2.1　步骤 5、研究试验设计

一旦确立了研究问题，就需要选择合适的研究设计，验证教育干预方法的有效性。全科教育干预研究中常用 5 种研究设计：单组后测、单组前测-后测、对照前测-后测、随机对照后测、随机对照前测-后测。教育研究与临床研究都应使用同样严谨的科研标准，但同时要考虑特定研究设计在教育研究中的优势与劣势，实际的研究设计选择需结合研究问题、研究群体及研究资源。我们将逐一阐述几种设计方法的优缺点并举例说明（见表 2）。

（1）**单组后测：**数据收集发生在教育干预结束时进行。该设计的主要优点是简单、经济、易操作，能快速获得培训反馈，因此在医学研究中比较普遍。因为只有一个数据收集点，学员投入的时间相对较少，数据容易分析[22]。然而，数据收集缺乏基线数据对比，因此结果也可能由干预以外的因素引起。

（2）**单组前测-后测：**在教育干预前后两个时间点收集数据，该方法是最常见的设计[4]。这种设计比单组后测设计更可靠，它可以更准确地检测干预前后的变化。其复杂度及花费适中，且无需对照组便能展示干预前后的变化。同时，缺点也十分明显，尽管在两个时间点收集了数据，但它不能准确地说明前后变化是否归因于干预或其他一些混杂影响。

（3）**对照前测-后测：**该设计是一种准实验方法，可以检测参与者前后变化是由于干预还是一些混杂因素的影响，并控制从第一次评估中学习的因素。与上述两种设计相比，对照前测-后测设计提供了更可靠的结果，且操作性强（比如一个班作为对照组，另一个班作为试验组），但研究前后的控制仍有一些局限性。需要确保两组的参与者特征的一致性，否则非测量因素的差异可能导致结果变化。此外，纳入对照组增加了数据收集和分析的量，从而增加了成本。由于对照组没有接受干预，可能存在失访。

（4）**随机对照后测：**参与者被随机分配到对照组和试验组，收集干预后的数据。随机对照试验能更有力地解释与干预相关的变化。课程中参与者的随机化意味着与选择或招募相关的偏差是最小的。该设计与随机对照前测-后测设计相比，需要较少投入，又能保持随机化。由于不能显示学习者的前后变化，需要基于检验假设的精确样本量以实现消除两组间的基线差异，但教育干预研究的样本量通常较少，因此在教育研究中并不常见[23]。

（5）**随机对照前测-后测：**参与者被随机分配到对照组和试验组，收集干预前后的数据。该设计是验证培训效果的金标准，外部效度高，可以通过随机化来控制混杂变量，最小化基线差异。但在教育环境中进行随机对照试验是困难的，因其研究所需的时间、费用、资源最多，难以确保参与者的依从性，且样本量要求高，往往需要多中心联合开展。

表 2 5 种教育干预研究设计方法的优缺点及在全科研究中的应用举例

Table 2 Advantages and disadvantages of 5 design methods in educational intervention research and their examples in general practice research

设计类型	优点	缺点	研究问题举例	研究设计举例
单组后测 X----O_1	简单、经济、易操作； 快速获得改进建议。	缺乏基线数据； 培训效果也许已经存在； 培训效果可能自然成熟引起； 培训效果可能是干预以外的因素引起。	工作坊教学模式是否可以提升全科医生的综合能力？[24]	研究对象：全科医生转岗培训学员； 干预措施：采用工作坊教学模式，包括全科门诊教学、病例讨论和全科教学查房； 评估方法：学员的现场问卷调查评价教学活动效果。
单组前测-后测 O_1--X--O_2	复杂度及花费适中； 无需对照组； 能展示干预前后的变化。	培训效果可能是干预以外的因素引起； 培训效果可能自然成熟引起； 培训效果可能来源于前后相同评估的学习。	基于社区实训评估基地对全科医生进行的培训是否有效果？[25]	研究对象：社区卫生中心的 30 名全科医生； 干预措施：基于社区实训评估基地对全科医生进行培训，包括理论知识、操作技能、体格检查以及医患沟通等。 评估方法：干预前后分别对全科医生各项能力水平进行评估比较。
对照前测—后测 E O_1-X-O_2 C O_1---O_2	对照干预以外的测量因素； 控制从第一次评估中学习； 比随机研究更具可行性。	复杂、需要大量资源； 两组参与者基线难达成一致。	痴呆症教育干预是否可以改善全科住院医生的态度，增加信心？[26]	研究对象：332 名全科住院医师； 干预措施：全科住院医师参加 3 小时的面对面讲习班； 评估方法：干预前后使用全科住院医生痴呆症信心和态度量表。
随机对照后测 E X---O_1 R C --O_1	控制混杂变量； 需要较少资源，又能随机化。	不能显示学习者的前后变化； 样本量要求高，两组参与者基线难达成一致。	对全科医生进行简短的非语言沟通培训是否可以提高患者对满意度？[27]	研究对象：16 名全科医生随机分为干预组和对照组； 干预措施：简短的非语言沟通培训，并观看和反思自己的接诊视频； 评估方法：干预后患者完成医疗访谈满意度量表。
随机对照前测—后测 E O_1-X-O_2 R C O_1---O_2	外部效度高； 最小化基线差异； 控制混杂变量。	最复杂，需要资源最多； 样本量大，往往需要多中心联合开展。	是否以学习者为中心的方法比传统方法更有效的提升全科医生沟通技巧？[28]	研究对象：荷兰 100 名全科医生，随机分为干预组和对照组； 干预措施：通过录像观察全科医生接诊中的沟通缺陷并反馈； 评估方法：干预前后通过对全科医生接诊视频评估沟通技能。

注：E：试验组；C：对照组；R：随机；X：干预；O_1：第一次测量；O_2：第二次测量。

2.2 **步骤 6、制定干预措施**

与临床研究类似，教育干预研究是复杂的，研究结果可能受到许多变量的影响，例如培训内容、教学环境、教学策略和培训对象。以医务人员高血压管理的医患沟通培训为例，不同研究采用不同的沟通培训内容、教学策略和时长，极大影响了教育干预对血压管理效果的异质性[29]。因此，制定教育干预措施需包含诸多教育相关的因素，如教学目标、教师资历、学员数量与知识储备、教学策略、培训日程、课程材料、培训环境、培训激励以及教学评价等。Meinema J G 等[30]基于 Phillips 等人开发的循证实践教育干预和教学报告指南(GREET)制定了一份教育干预措施的核查表，分 3 个部分和 17 条准则，全面描述医学教育中的教育干预措施，可以有效协助研究者制定更详细的教育干预措施，以提升研究的透明度和可复制性，更好地解释研究结果。

2.3 **步骤 7、干预结果评估**

教育干预研究需要选择合适的结果指标以评估并证实培训的有效性。在众多教育评估模型中，柯氏四级评估模型 （Kirkpatrick's four-level model of evaluation）在教育研究领域应用最广，并被最佳证据医学教育（Best Evidence Medical Education）协作组织作为评价各种教育模式或方法的“证据”分级标准[31, 32]。柯氏四级评估模型不仅关注医学教育干预对学习者的满意度、知识和技能改善，也评估培训对学习者工作行为和患者医疗结局的影响，其主要包含以下 4 个层次（见表 3）：反应层(满意度)、学习层(态度、知识和技能)、行为层（工作行为改变）及成果层（患者影响）[33]。Yardley 等在回顾医学教育干预研究相关文献发现，24%的研究结果评估处于第 1 级，64%处于 2 级，只有 12%处于 3 级和 4 级[31]。但在研究设计过程中，并非一味追求更高层级的结果评估，而是需要结合研究目的和资源综合考量。总体上讲，结果评估层次较低但干预方法高度创新的研究更容易引人关注，而创新水平较低的研究通常需要更高层次的结果评估。

在决定结果评估层次后，接下来就需要选择合适的评估方法和工具，比如通常采用问卷评估学习者对培训的满意度，知识测验评估学习者的知识掌握程度，标准化病人评估技能改善，以及病历质控评估实际行为改变（见表 3）。不同的结果评估层次有不同的评估方法，而相同的评估方法又有不同的评估工具，比如在培训医患沟通技能后评估全科医生门诊中的行为改变，可以通过自我报告、研究者直接观察、录像、匿名标准化病人等不同的方法进行评估，而其中研究者直接观察可以采用自制打分表、SEGUE 量表，卡尔加里剑桥指南等评估工具[34]。著名的医学教育学者 Cook 教授建议先确定结果评估层次，然后是评估方法，最后是评估工具[35]。每种评估方法都有其优缺点，而评估工具间有着不同的信效度[36]。研究者需尽力选择已经存在且信效度高的评估工具，从而避免在开发评估工具上花太多时间。如果没有现成的评估工具，可以采用折中的方法，结合已有工具和研究需求，开发新的评估工具。建议通过预实验验证评估工具，以发现在数据收集过程中存在的问题[35]。

表 3 **柯氏四级评估模型及应用举例**

Table 3 Examples of Kirkpatrick's four-level evaluation model

级别	评估类型	具体内容	常用评价方法	举例（全科医生结肠癌筛查培训）
第 1 级	反应层(Reaction)	学员对培训的评价：参与度、满意度、接受度、实用性	学员问卷 学员访谈	全科医生对“如何在全科门诊中开展结肠癌筛查”课程的满意度、实用性的问卷调查
第 2 级	学习层(Learning)	学员在培训中的收获：知识、技能、态度、信心、意愿	知识测验 标准化病人 问卷量表	1、知识测验：测量全科医生对结肠癌的相关知识的掌握情况； 2、标准化病人：评估全科医生对社区居民开展结肠癌筛查的沟通及决策的技能掌握情况 3、问卷量表：评估全科医生对社区居民开展结肠癌筛查的意愿和态度
第 3 级	行为层(Behavior)	培训后学员在实践中的行为改变	录像或直接观察 匿名标准化病人 病历查阅	1、直接观察：研究人员通过录像观察全科医生实际接诊中是否开展结肠癌筛查； 2、病历查阅：符合筛查的患者是否筛查结肠癌

第 4 级	成果层 (Result)	培训带来的机构改变或患者获益	患者问卷 实验指标 机构统计报告	1、培训后 1 月机构开展结肠癌筛查的数量； 2、培训后 1 年内结肠癌患者的诊断数量； 3、培训后管辖社区结肠癌的患者死亡数量

3　阶段三：研究实施

3.1　步骤 8、组建研究团队

良好的研究团队对课题的成功实施至关重要。主要研究者应确保团队里有合适的人员，不仅应具备研究相关的专业背景和经验，也需要对研究的热情和投入。经验较少但愿意奉献的低年资成员可能与具有许多研究经验但付出有限的资深成员一样有价值。因此，在一个全科教育研究团队里应主要包含两类成员，一类是课题的执行者，主要开展教育的评估、干预以及数据的采集，可以邀请与您平级的同事，或低年资的医师、研究生、规培生、实习生加入其中；另一类是课题的专家顾问，比如资深的教育研究者、统计学专家，丰富教龄的教师。这些顾问专家可能不是临床医生，不熟悉具体的临床环境，但他们在教育理论和研究方法上具有丰富的知识和经验，可以为教育课题的高质量开展提供绝佳的建议[22]。如果在您的机构或医院没有这样的人员，可以尝试联系对您研究感兴趣的外机构且有相关经验的合作者[37]。

作为主要研究者，一方面需要通过开放的讨论，让成员知晓自己在教育研究实施过程中所扮演的角色和期待；另一方面，也要尽力思考如何让团队成员在研究中获益，比如共同发表文章、学习新的科研技能、结识新的朋友、获得学位、增加职业履历、劳务报酬以及获得您的导师推荐等，每个成员的期待都有所不同，需要研究者仔细的权衡，才能有效地凝聚团队成员并做出持续的贡献。一个有效的策略是在研究项目的早期阶段就达成对作者排序和贡献的共识，甚至签署合作的协议，以避免后期团队成员之间的冲突[38]。

3.2　步骤 9、获取研究资源

研究团队需要仔细考量并争取教育研究项目实施过程中所需要的各种资源。首先，尝试说服机构或相关负责人认可并支持你的教育干预研究项目，让利益相关者认识到该教学研究可能为机构和学员带来的潜在价值，从而为你在教育干预中的学员组织、场所、设备等提供便利。其次，大多数忙碌的全科临床师资有着自己的临床工作或教学任务，几乎很难抽出足够的时间开展教育研究。克服这一困难的建议是让教育研究围绕已经运行的全科课程或培训项目，这会减少许多资源的投入。最后，有必要为您的教育干预项目申请资金，许多学校或医疗机构会为教学创新改革项目提供内部的资金支持，研究者可以充分利用这样的资源去开展教育干预研究的预实验，初始研究成果可以为下一步申请外部基金提供有力的支持。目前，国内许多城市或省份的健康卫生委员会或全科专业协会均可为全科教育干预项目提供资金支持，也有专门针对全科教育研究的基金，如全国医学专业教育指导委员会与中国医师协会联合启动的全科医学教育教学研究[39]。如果读者不熟悉如何撰写教育研究的申请标书，可以参见相应的标书撰写指南[40, 41]。

3.3　步骤 10、获得伦理审批

尽管研究者普遍认同在临床研究中需要获得伦理审批，在医学教育研究中却差异很大。一方面，研究者可能认为教育研究不会直接纳入患者，也倾向认为教育干预措施会给学生带来益处，因此容易忽视其潜在的风险；另一方面，教育研究的双重目的（教育与研究）也会掩盖研究中可能存在的伦理问题[42]。Hally E 等的调查显示在四本主要的医学教育期刊发表的原始研究，仅 5%经过了机构伦理委员会的审批[43]。而且，学生在教育研究中仍是“弱势群体”，由于研究人员同时具有教师的角色，学生可能会担心研究的配合程度会影响教师对自己的评价或成绩，而被迫参与研究；研究也会占用学生本可以用于临床实践、学习其他科目或休闲娱乐的时间；甚至在研究过程中的不良表现会影响学生在研究以外的自信心；同时，研究也存在暴露学生隐私的风险[42]。因此，和临床研究一样，医学教育干预研究仍需要在保证真实性、有效性的基础上确保学生的自主选择、风险最小化和社会公正，在开展研究之前经过机构伦理委员会的审批。尽管大多数医学教育研究可以豁免伦理审批，但这个决定是由伦理委员会做出，而非研究者本身。目前医学教育的杂志也逐渐变得更加严格，要求在发表前提供机构伦理委员会的批准[44]。

研究者在设计研究时可以采取一系列措施将伦理风险减少到最小，从而顺利获得伦理委员会的认可。首先，研究计划书要清晰地描述知情同意的过程，详细说明学生可以自由选择参与研究，而不用担心拒绝会带来任何后果，并且在同意后仍可以选择退出[45]。其次，在调查评估时尽量采取匿名的方式，或者使用代码以减少学生身份被识别的风险，研究者也需要妥善保存研究数据，避免泄露学生的隐私[45]。同时，出于公平和公正的考虑，伦理委员会更倾向认可交叉设计的研究方案，而非仅试验组学生获得相应的教育干预[46]。学者 Egan-Lee 和学者 Boileau E 分别总结了医学教育研究获得伦理审批的 12 项技巧[45]和 12 个可避免的陷井[47]，供医学教育研究的新手进一步学习和参考。最后，研究者需要理解伦理委员会在研究中所发挥的作用，通过合作的态度与伦理委员会专家保持沟通，共同确保教育研究工作能安全并符合伦理地开展。

3.4 步骤 11、项目实施

成功的教学研究项目不仅需要精心、严谨的设计，更需要务实、严格地执行。当研究项目真正进入实施阶段时，项目负责人须转变角色，从研究者转变为项目的“管理者”。除了组建研究团队和获取研究资源，作为项目管理者，需要在实施过程中制作一份详尽的时间安排，明确各项工作的起止时间，包括招募受试者、课程培训、收集数据的节点，数据录入与分析以及发表论文的时间，并明确各团队成员在各个环节中的责任。推荐使用甘特图（Gantt Chart），有助于研究负责人跟进实施进度。但是，课题负责人要对项目实施的顺利程度抱有合理的预期，总会遇到各种困难和障碍，如部分学员不配合调查，疫情对培训进度影响，干预组与对照组之间的沾染，无法按照预期时间开展结果评价等。这时在团队成员间建立固定的例会制度（如每周 1 次），定期汇报和沟通研究的进展和遇到的难题，通过头脑风暴、开放式地讨论等方式，共同商议和解决课题实施中的问题。推荐初学者能提前熟悉项目管理中的基本知识和技巧[48]从而更好地推进研究的开展。

4 阶段四：发表与评价

4.1 步骤 12、文章撰写和发表

教育干预研究报告的成功发表不仅有利于研究者个人的学术发展，扩大自身的学术影响力，也有助于教育干预措施的转化和推广，为其他研究人员提供参考。然而，医学教育干预研究的报告质量普遍较差，经常缺失许多科学报告的基本要素[5]。Meinema JG 等在 2019 年的系统综述发现，在纳入的 105 篇医学教育干预研究中，52.4%的研究未陈述研究的认知框架，43.8%未描述学习培训目标，53.3%缺乏学生满意度的报告，报告质量平均分仅为 15.9（总分 32 分），76%的文章仅在 11-20 分之间[30]。因此，了解教育研究报告的撰写规则对成功发表十分重要，以确保研究遵循相关的国际公认准则。目前已开发有教育干预研究的学术报告规范（见表 4），如教育干预研究的 BMJ 指南[49]，Cook 报告的关键元素[5]，基于循证实践的教育干预和培训报告指南（GREET）[30, 50]等。以上报告规范提供了如何撰写教育干预文章的详尽说明，可以为全科教育研究者节约时间，减少撰写中遇到的盲区，提升报告的质量，也推荐在研究设计阶段能参考上述报告规范。

此外，选择合适的期刊也是成功发表的重要策略。有关全科教育干预研究的发表杂志，一方面可以考虑全科医学的中英文杂志，也可以考虑教育类的中英文杂志。笔者以中国知网数据库为数据来源检索已经发表的中文的全科教育研究，以“(标题=全科) OR (标题=社区医生) OR (标题=全科规范化)OR (标题=家庭医生)OR (标题=乡村医生)”和“(标题=教育) OR (标题=培训) OR (标题=课程)OR (标题=教学)”进行检索，检索时间均设定为 2010 年至今。英文文献检索以 Pubmed 数据库为数据来源，检索式为：(((((education[Title]) OR (curriculum[Title])) OR (teaching[Title])) OR (training[Title])) OR (course[Title]) AND ((((((family physician[Title])) OR (general practitioner[Title])) OR (family practice[Title])) OR (general practice[Title])) OR (primary care physician[Title])) OR (GP residents[Title])，列出了全科教育研究发表前 5 位的中英文杂志（见表 5），供读者在未来投稿时参考。

表 4:医学教育干预研究的报告规范总结

Table 4 Summary of reporting norms for medical education intervention studies

名称	出版时间	作者	研究类型	评估内容简介
BMJ guidelines[49]	1999	Luisa Dillner 等	教育干预研究	指南分为 4 个部分（概述、理论考虑、培训介绍和设计、讨论）和 18 个项目。

Cook Key elements of reporting [5]	2007	Cook DA 等	教育干预研究	基于指南的质量报告特征包括：文献综述、概念框架、研究意图、研究设计、干预组和对照组，以及伦理考虑。
GREET [50]	2014	Phillips AC 等	教育干预研究	确定了 39 个项目来描述循证实践的教育干预，提供仅用于描述干预措施的详细信息，与研究设计的现有报告指南结合使用。
Recommendations for Reporting Mastery Education Research in Medicine (ReMERM) [51]	2015	Cohen ER 等	掌握性学习研究	指南包含 22 个类别和 38 个基本项目，分为六个部分:标题和摘要、简介、方法、结果、讨论和其他信息。
The improvement of the GREET[30]	2019	Meinema JG 等	教育干预研究	确定了 GREET 未涵盖的其他项目，在清单中添加了以下标准：学习需求、干预措施发展过程、背景和环境、参与者以及评估和满意度。

GREET：Guideline for Reporting Evidence-Based Practice Educational Interventions and Teaching

表 5 全科教育研究常发表的中英文杂志

Table 5 Common published English and Chinese journal for educational research in General practice

排序	中文杂志	构成比（%）	类别	英文杂志	构成比（%）	类别
1	中国全科医学	8.11	全科	Education for primary care	17.89	全科
2	继续医学教育	3.95	教育	BMC medical education	5.99	教育
3	中华全科医学	3.17	全科	The British journal of general practice	5.99	全科
4	卫生职业教育	2.77	教育	Australian family physician	5.17	全科
5	中国高等医学教育	2.56	教育	The medical journal of Austrilia	2.89	综合

4.2 步骤 13、反思与评价

尽管精心的设计，由于客观条件和资源的限制，实际开展的研究总会存在各种的瑕疵，“样本代表性差、描述不完整、数据处理不当”等问题都将直接造成研究质量下降，从而导致研究结果存在偏倚，缺乏科学价值，[52]，因此研究结束后的反思对未来教学研究的开展显得尤为重要。推荐应用科学规则对其质量进行评价（Critical Appraisal），从研究对象的代表性、干预描述、偏倚控制、结果评估和分析报告方法五个领域进行[53]。

目前已出版的质量评价工具中，有多个适用于医学教育研究，基于不同的评价内容和标准，其侧重点有一定差异（见表 6）。这些评价工具为教育研究项目提供了一个规划和评估的框架，为研究人员在项目设计阶段、实施阶段、数据收集分析和结果报告等方面提供高质量的参考标准；同时，也有助于学术期刊审稿人、同行评议、读者们对教育干预研究的报道进行同质化的审核[54]。

表 6 医学教育干预研究相关的质量评价工具

Table 6 Quality evaluation tools for medical educational interventional research

质量评价工具	作者与出版时间	评价内容	适合研究类型
医学教育研究质量工具（MERSQI）[55]	Reed DA 2007	研究设计、样本、数据类型、评价工具效度、数据分析、评价结局。	实验性、类实验性、观察性研究
医学教育最佳证据(BEME)[56]	Littlewood S 2005	基于柯氏评估模型对研究的结果和结论进行评价，最终归类为六个级别的证据。	基于临床或社区环境的教学研究
纽卡斯尔-渥太华教育文献质量评价量表(NOS-E) [57]	Cook DA 2008	基于纽卡斯尔-渥太华质量评定量表对教育研究文献的样本代表性、可比性、研究设计、结局评估方法进行质量评价。	非随机对照研究、队列研究、互联网干预性研究
急诊学术医学教育研究评分标准(AEM)[58]	Farrell 2011	引言、度量、数据收集、数据分析、讨论、局限性、创新性、概括性、写作清晰性	定量研究中教育干预的假设检验调查和测量

MERSQI:Medical Education Research Study Quality Instrument;BEME:Best Evidence in Medical Education;NOS-E:Newcastle–Ottawa Scale Education; AEM:Education Research Scoring Metrics

5 讨论

由于中国独特的卫生服务体系和社会环境，全科教育者要意识到全科医生的培养将会面临许多新的问题和挑战，需要通过严谨且有效的方式去探索并改变。我们试图为全科教育者开展干预研究提供一个初步的思路和步骤，以指导未来的全科教学实践，提高研究的严谨性，并将创新的教学实践发表到期刊，分享给更多的全科教育者学习和借鉴。

在开展教育干预研究之前，全科教育者需要学会规范地设计医学教育课程，如需求评估、教学内容、课程目标、教育策略、实施课程、评估课程等，这是教学研究的前提，可以参看 SCHNEIDERHAN 等的著作以及阅读书籍《医学教育课程开发：六步法》[59]。在此基础上融合上述的研究步骤，将教学项目转化为教学科研[60]。对于初学者，我们建议能与机构内外经验丰富的全科教育者和研究者合作，积累研究经验，确保研究的创新性与严谨性。如果研究资源有限，可以在已经运行的全科培训项目的基础上开展教学研究，以提高研究的可行性。

由于篇幅有限，本文无法深入探讨每一个研究步骤，也不能囊括开展教育干预研究中所有步骤和方法，如样本量计算、数据采集以及数据分析等[9]。读者可以阅读本文每一个步骤中所引用的参考文献，也推荐阅读书籍《医学教育研究概论》[19]、《教育研究：定量、定性和混合方法》进一步深入学习。中国大学 MOOC 有教育研究方面的公开课程供读者学习，如西南大学和浙江大学《教育研究方法》[61]虽然方法、书籍和课程有助于建立教育干预研究的基础知识，但掌握这种研究方法的最佳策略是不断地实践和总结。同时，希望读者谨记的是，教育干预研究也有其局限性，并不能回答所有的教育研究问题，如描述性问题采用问卷调查，探索性问题需采用定性研究，需根据不同的研究问题制定不同的研究方法[62]。而且即使教育干预研究经过良好的设计，能较好的提高其内部效度，由于教育干预的效果和当地的背景（学习者、教师、文化等）密切相关，研究结果的外推性受到很大的限制，因此融入理论/概念框架显得尤为重要。

6 结语

作为全科教育者，一方面我们有责任让全科学员满意，并提升学员的理论知识和技能水平，培养有胜任力的全科医生。另一方面，也希望培训措施能带来实际的效果，增进病人的健康，这也是开展全科培训的初衷。希望通过本文所呈现的方法和步骤，如构建研究问题，设计试验方法和结果评估，借助教育研究的报告规范和质量评价工具等，全科教育者能够科学、规范、有效地开展干预研究，在全科教育理论与临床实践之间发挥重要的桥梁作用。

作者贡献：邹川、林恺、曾欣负责文章的构思与设计；邹川、林恺负责文章的可行性分析、论文修订；欧健铭、彭涛、周燕、肖纯瑶、陶洪霞负责文献/资料收集与整理、论文撰写；曾欣负责英文修订；陈庆奇（学术全科医生）负责审查和编辑；林恺、邹川负责文章的质量控制及审校，并对文章整体负责，监督管理。

利益冲突：本文无利益冲突。

志谢：感谢南方医科大学-美国国际医学教育与研究促进基金会中国区域中心（SMU-FAIMER Regional Institution）对团队成员所开展的教育研究指导与培训。

参考文献：

[1] Zou C, Liao X Y, Spicer J, et al. Ten years' GP training in China: progress and challenges [J]. Br J Gen Pract, 2020, 70(699): 511-512.

[2] Harden R M, Grant J Fau- Buckley G, Buckley G Fau- Hart I R, et al. Best Evidence Medical Education [J]. Adv Health Sci Educ Theory，Pract, 2000, 1573-1677 (Electronic)):

[3] Cruz-Cunha M M, Miranda I M, Gonçalves P. What is Educational Intervention [J]. 2013,

[4] Cook D A, Beckman T J, Bordage G. Quality of reporting of experimental studies in medical education: a systematic review [J]. Med Educ, 2007, 41(8): 737-745.

[5] Cook D A, Beckman T J, Bordage G. Quality of reporting of experimental studies in medical education: a systematic review [J]. Medical Education, 2007, 41(8):

[6] Yarris L M, Deiorio N M. Education research: a primer for educators in emergency medicine [J]. Acad Emerg Med, 2011, 18 Suppl 2(S27-35.

[7] Boet S, Sharma S, Goldman J, et al. Review article: medical education research: an overview of methods [J]. Canadian journal of anaesthesia = Journal canadien d'anesthésie, 2012, 59(2): 159-170.

[8] Cleland J, Durning S J. Researching Medical Education [M]. Researching Medical Education, 2015.

[9] Tavakol M, Sandars J. Quantitative and qualitative methods in medical education research: AMEE Guide No 90: Part I [J]. Med Teach, 2014, 36(9): 746-756.

[10] Mann K V. The role of educational theory in continuing medical education: has it helped us? [J]. J Contin Educ Health Prof, 2004, 24 Suppl 1(S22-30.

[11] Brown J, Bearman M, Kirby C, et al. Theory, a lost character? As presented in general practice education research papers [J]. Med Educ, 2019, 53(5): 443-457.

[12] Meinema J G, Buwalda N Fau- Van Etten-Jamaludin F S, Van Etten-Jamaludin Fs Fau- Visser M R M, et al. Intervention Descriptions in Medical Education: What Can Be Improved? A Systematic Review and Checklist [J]. Acad Med, 2019, 1938-808X (Electronic)):

[13] Bordage G. Conceptual frameworks to illuminate and magnify [J]. Med Educ, 2009, 43(4): 312-319.

[14] Abramson E L, Paul C R, Petershack J, et al. Conducting Quantitative Medical Education Research: From Design to Dissemination [J]. Acad Pediatr, 2018,

[15] Slotnick H B, Shershneva M B. Use of theory to interpret elements of change [J]. Journal of Continuing Education in the Health Professions, 2010, 22(4):-.

[16] Shaughnessy A F, Duggan A P. Family medicine residents' reactions to introducing a reflective exercise into training [J]. Educ Health (Abingdon), 2013, 26(3): 141-146.

[17] Malaty J, Williams M, Carek P J. Impact of Providing Data on Family Medicine Practice Management Education [J]. Family Medicine, 2020, 52(6): 432-434.

[18] Klein D, Schipper S. Family medicine curriculum: improving the quality of academic sessions [J]. Can Fam Physician, 2008, 54(2): 214-218.

[19] 王维民. 医学教育研究概论 [M]. 北京：北京大学医学出版社, 2020.

[20] Riva J J, Malik K M, Burnie S J, et al. What is your research question? An introduction to the PICOT format for clinicians [J]. J Can Chiropr Assoc, 2012, 56(3): 167-171.

[21] Hulley S B, Cummings S R, Browner W S. Conceiving the research question and developing the study plan [M]. Designing Clinical Research 4th ed. Philadelphia; Lippincott Williams and Wilkins. 2013: 14-22.

[22] Boet S, Sharma S, Goldman J, et al. Review article: medical education research: an overview of methods [J]. Can J Anaesth, 2012, 59(2): 159-170.

[23] Mcgaghie W C, Issenberg S B, Cohen E R, et al. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence [J]. Acad Med, 2011, 86(6): 706-711.

[24] 魏娜娜, 梁冰, 郦忆文, et al. 工作坊教学模式在全科医生转岗培训中的实践探索 [J]. 中华全科医学, 2021, 19(10): 4.

[25] 左文杰, 段晚霞, 谷英英, et al. 基于社区实训评估基地的全科医生培训效果研究 [J]. 中外医学研究, 2016, 14(17): 2.
[26] Mason R, Doherty K, Eccleston C, et al. Effect of a dementia education intervention on the confidence and attitudes of general practitioners in Australia: a pretest post-test study [J]. BMJ Open, 2020, 10(1): e033218.
[27] Little P, White P, Kelly J, et al. Randomised controlled trial of a brief intervention targeting predominantly non-verbal communication in general practice consultations [J]. Br J Gen Pract, 2015, 65(635): e351-356.
[28] Hobma S, Ram P, Muijtjens A, et al. Effective improvement of doctor-patient communication: a randomised controlled trial [J]. Br J Gen Pract, 2006, 56(529): 580-586.
[29] Yao M, Zhou X Y, Xu Z J, et al. The impact of training healthcare professionals' communication skills on the clinical care of diabetes and hypertension: a systematic review and meta-analysis [J]. BMC Fam Pract, 2021, 22(1): 152.
[30] Meinema J G, Buwalda N, Etten-Jamaludin F V, et al. Intervention Descriptions in Medical Education: What Can Be Improved? A Systematic Review and Checklist [J]. Academic medicine: journal of the Association of American Medical Colleges, 2019,
[31] Yardley S, Dornan T. Kirkpatrick's levels and education 'evidence' [J]. Medical Education, 2011, 46(1):
[32] Reeves S, Fletcher S, Barr H, et al. A BEME systematic review of the effects of interprofessional education: BEME Guide No. 39 [J]. Med Teach, 2016, 38(7): 656-668.
[33] Kirkpatrick D L. Evaluating training programs: Inside the four levels [J]. Mcgraw Hill Professional, 2008,
[34] 邓黎黎, 廖晓阳, 伍佳, et al. 国外医患沟通模式对我国全科医生沟通技能培训的启示 [J]. 中国全科医学, 2021, 24(13): 6.
[35] Cook D A. Twelve tips for evaluating educational programs [J]. Med Teach, 2010, 32(4): 296-301.
[36] Anderson T R, Rogan J M. Bridging the educational research-teaching practice gap: Tools for evaluating the quality of assessment instruments [J]. Biochem Mol Biol Educ, 2010, 38(1): 51-57.
[37] Scott K, Caldwell P, Schuwirth L. Ten steps to conducting health professional education research [J]. Clinical Teacher, 2015, 12(4): 272-276.
[38] Sommers M S. Negotiating journal authorship: strategies and hazards [J]. Clin Nurs Res, 2011, 20(2): 115-119.
[39] 医专业学位委. 关于公布 2021 年全国全科医学教育教学研究课题立项结果的通知[EB/OL] (2022-01-08)[2022-03-20]. http://www.medgrad.cn/site/content/2824.html.
[40] Blanco M A, Lee M Y. Twelve tips for writing educational research grant proposals [J]. Med Teach, 2012, 34(6): 450-453.
[41] Blanco M A, Gruppen L D, Artino A R, Jr., et al. How to write an educational research grant: AMEE Guide No. 101 [J]. Med Teach, 2016, 38(2): 113-122.
[42] Keune J D, Brunsvold M E, Hohmann E, et al. The ethics of conducting graduate medical education research on residents [J]. Academic Medicine, 2013, 88(4): 449-453.
[43] Hally E, Walsh K. Research ethics and medical education [J]. Medical teacher, 2016, 38(1): 105-106.
[44] Eva K W. Research ethics requirements for Medical Education [J]. Medical Education, 2010, 43(3): 194-195.
[45] Egan-Lee E, Freitag S, Leblanc V, et al. Twelve tips for ethical approval for research in health professions education [J]. Med Teach, 2011, 33(4): 268-272.
[46] Kraus C K, Guth T, Richardson D, et al. Ethical considerations in education research in emergency medicine [J]. Acad Emerg Med, 2012, 19(12): 1328-1332.
[47] Boileau E, Patenaude J, St-Onge C. Twelve tips to avoid ethical pitfalls when recruiting students as subjects in medical education research [J]. Med Teach, 2018, 40(1): 20-25.
[48] （美）项目管理协会. 项目管理知识体系指南 PMBOK 指南 第 6 版 [M]. 北京: 电子工业出版社, 2018.
[49] Guidelines for evaluating papers on educational interventions [J]. BMJ, 1999, 318(7193): 1265-1267.
[50] Phillips A C, Lewis L K, Mcevoy M P, et al. A Delphi survey to determine how educational interventions for evidence-based practice should be reported: stage 2 of the development of a reporting guideline [J]. BMC Med Educ, 2014, 14(159.
[51] Cohen E R, Mcgaghie W C, Wayne D B, et al. Recommendations for Reporting Mastery Education Research in Medicine (ReMERM) [J]. Acad Med, 2015, 90(11): 1509-1514.
[52] Jennifer. P, Barton. B. Medical statistics: A guide to data analysis and critical appraisal [M]. Oxford: Blackwell Publishing;BMJ Books, 2005.

[53] Price E G, Beach M C, Gary T L, et al. A systematic review of the methodological rigor of studies evaluating cultural competence training of health professionals [J]. Acad Med, 2005, 80(6): 578-586.
[54] Cook D A, Reed D A. Appraising the quality of medical education research methods: the Medical Education Research Study Quality Instrument and the Newcastle-Ottawa Scale-Education [J]. Acad Med, 2015, 90(8): 1067-1076.
[55] Reed D A, Cook D A, Beckman T J, et al. Association Between Funding and Quality of Published Medical Education Research [J]. Jama the Journal of the American Medical Association, 2007, 298(9): 1002-1009.
[56] Littlewood S, Ypinazar V, Margolis S, et al. Early practical experience and the social responsiveness of clinical education: Systematic review [J]. BMJ (Clinical research ed), 2005, 331(387-391.
[57] Cook D A, Levinson A J, Garside S, et al. Internet-Based Learning in the Health Professions: A Meta-Analysis [J]. Jama the Journal of the American Medical Association, 2008, 300(10): 1181-1196.
[58] Farrell S E, Kuhn G J, Coates W C, et al. Critical appraisal of emergency medicine education research: the best publications of 2013 [J]. Acad Emerg Med, 2014, 21(11): 1274-1283.
[59] Thomas P A, 周玉皆. 医学教育课程开发 六步法 原书第 3 版 [M]. 北京：科学出版社, 2019.
[60] Zou C, Liao X. Transforming a medical curriculum development into an education research [J]. Family Medicine and Community Health, 2019, 7(3):
[61] 陈时见, 周琴, 王正青. 教育研究方法[EB/OL] (2022-03-08)[2022-07-01]. https://www.icourse163.org/course/ZJU-1206404806?from=searchPage.
[62] Ringsted C, Hodges B, Scherpbier A. 'The research compass': an introduction to research in medical education: AMEE Guide no. 56 [J]. Med Teach, 2011, 33(9): 695-709.